SCIENCE CHINA
Life Sciences

SPECIAL TOPIC: Biosynthesis and regulation of secondary metabolites in microorganisms
• RESEARCH PAPER •

July 2013 Vol.56 No.7: 619–627
doi: 10.1007/s11427-013-4504-2

Mining of a streptothricin gene cluster from *Streptomyces* sp. TP-A0356 genome via heterologous expression

LI JinE[1], GUO ZhengYan[1], HUANG Wei[2], MENG XiangXi[1,3], AI GuoMin[1], TANG GongLi[2*] & CHEN YiHua[1*]

[1]*State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;*
[2]*State Key Laboratory of Bio-organic and Natural Product Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China;*
[3] *University of Chinese Academy of Sciences, Beijing 100049, China*

Received April 26, 2013; accepted May 28, 2013

Streptothricins (STs) are used commercially to treat bacterial and fungal diseases in agriculture. Mining of the sequenced microbial genomes uncovered two cryptic ST clusters from *Streptomyces* sp. C and *Streptomyces* sp. TP-A0356. The ST cluster from *S.* sp. TP-A0356 was verified by successful heterologous expression in *Streptomyces coelicolor* M145. Two new ST analogs were produced together with streptothricin F and streptothricin D in the heterologous host. The ST cluster was further confirmed by inactivation of gene *stnO*, which was proposed encoding an aminomutase supplying β-lysines for the poly-β-Lys chain formation. A putative biosynthetic pathway for STs is proposed based on bioinformatics analyses of the ST genes and experimental evidence.

genome mining, streptothricin, *Streptomyces*, heterologous expression, biosynthesis

Citation: Li J E, Guo Z Y, Huang W, et al. Mining of a streptothricin gene cluster from *Streptomyces* sp. TP-A0356 genome via heterologous expression. Sci China Life Sci, 2013, 56: 619−627, doi: 10.1007/s11427-013-4504-2

As members of the earliest discovered antibiotics from actinomycetes, streptothricins (STs) represent a group of natural products containing the unique streptolidine base such as nourseothricins, albothricin and LL-AC541 [1–4] (Figure 1A). Besides the streptolidine lactam moiety, STs are composed of a carbamoylated glucosamine conjugated with a poly-β-Lys chain, which can elongate up to seven β-lysines. STs can inhibit protein synthesis and show antibacterial, antifungal, insecticidal and phytogrowth inhibition bioactivities [5]. However, the severe cytotoxicity of STs impedes their clinical use [6]. In agriculture, STs are used to prevent bacterial and fungal diseases of crops, fruits and vegetables in China. In addition, nourseothricins and its resistant gene encoding nourseothricin acetyl transferase have been developed as a marker system used in molecular biological researches [7].

Biosynthesis of STs has been investigated initially by isotope feeding experiments. It has been established that the β-Lys residues in ST poly-β-Lys chain are derived from L-Lys [8]. Two incomplete ST clusters were cloned from *Streptomyces rochei* F20 and nourseothricin producer *Streptomyces noursei* [9,10]. Preliminary studies on those two clusters suggested that a non-ribosomal peptide synthetase (NRPS) machinery was involved in the poly-β-Lys chain biosynthesis. Recently, Maruyama et al. [11] cloned two complete ST clusters from *S. rochei* NBRC12908 and *Streptomyces lavendulae* NBRC12789, and proposed a detailed NRPS model for the poly-β-Lys chain biosynthesis.

*Corresponding author (email: chenyihua@im.ac.cn; gltang@sioc.ac.cn)

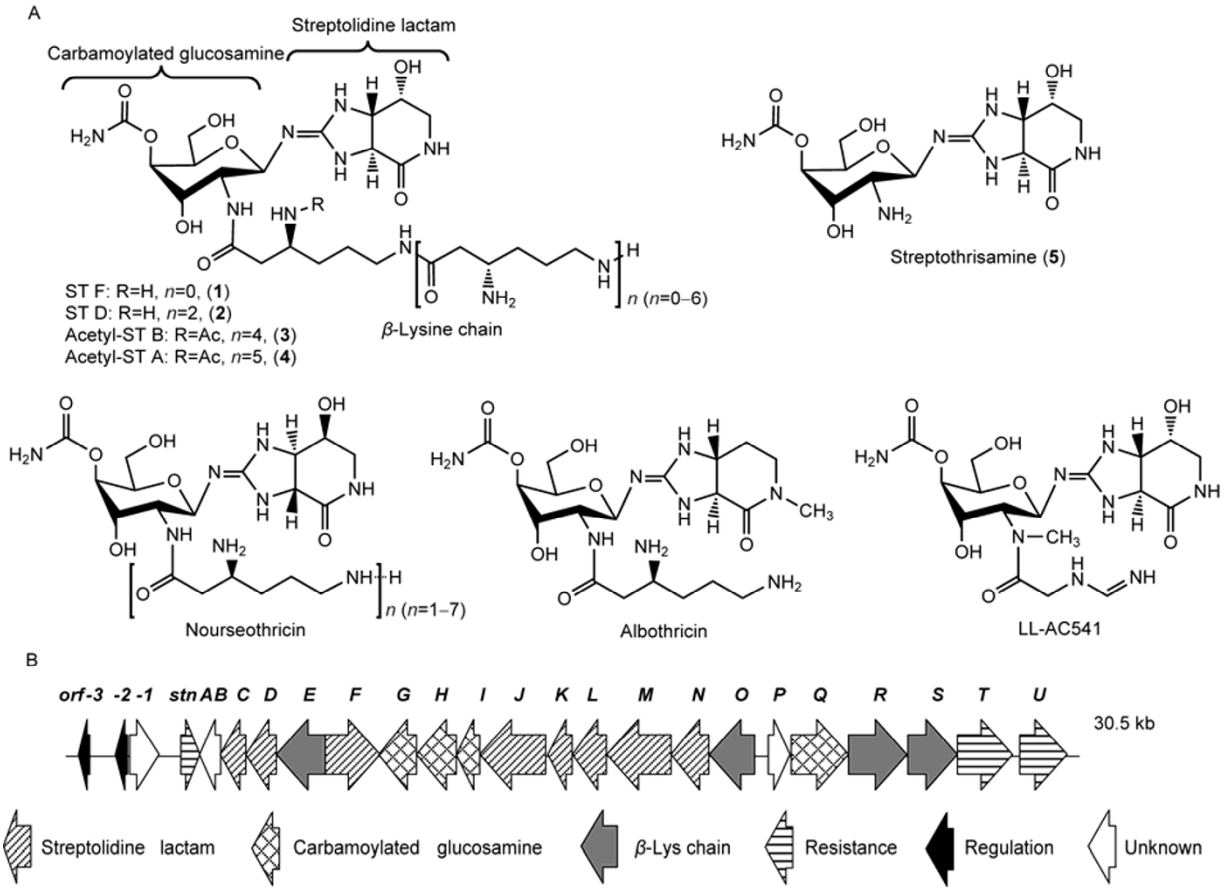

Figure 1 Structures of diverse streptolidine containing compounds and organization of the streptothricin cluster from *S.* sp. TP-A0356. A, Structures of streptothricins, streptothrisamine and other streptolidine containing compounds. B, Organization of the streptothricin biosynthetic gene cluster from *S.* sp. TP-A0356.

In *S. rochei* NBRC12908, three enzymes are involved in the poly-β-Lys chain formation. A stand-alone adenylation (A) domain (Orf5) loads one β-Lys onto the peptidyl carrier protein (PCP) moiety of the didomain protein Orf18 (PCP-C, C for condensation), another A domain (Orf19) elongates the PCP tethered poly-β-Lys chain by forming ε-amide bond between β-Lys iteratively, the C domain of Orf18 then releases the poly-β-Lys chain by condensing it with streptothrisamine to generate STs. Isotope labeling results showed that the unique streptolidine moiety is derived from L-Arg, and the carbamoylated glucosamine is from glucosamine [12,13]. So far, little is known about the biosynthetic mechanisms of these two parts.

STs can be isolated from soil dwelling actinomycetes at a 10^{-1} frequency in antibacterial agents screening [14]. Considering the wide distribution of STs in nature, we believe that ST clusters should exist in the sequenced genomes. Indeed, two putative ST biosynthetic gene clusters were obtained from *Streptomyces* sp. TP-A0356 (Figure 1B) and *Streptomyces* sp. C genomes when we searched the open accessed databases and several sequenced *Streptomyces* genomes in our lab using ST biosynthetic genes as probes. Although there are still some gaps in the *S.* sp. C ST cluster, it is clear that both of the cryptic ST clusters are very similar to the *S. rochei* NBRC12908 ST cluster (Figure S1 in Supporting Information) [11].

In this paper, we report the identification of the ST cluster from *S.* sp. TP-A0356 via heterologous expression. Two new ST analogs were obtained together with ST-F and ST-D in the heterologous host. The ST cluster was further confirmed by gene inactivation. A biosynthetic pathway of STs was proposed based on bioinformatics analysis and experimental results.

1 Materials and methods

1.1 Bacterial strains, media and plasmids

Escherichia coli JM109 was used to propagate plasmids. *E. coli* ET12567/pUZ8002 and BW25113/pKD20 were used for intergeneric conjugation and λ RED-mediated PCR-targeting mutagenesis respectively [15,16]. *Streptomyces coelicolor* M145, *Streptomyces albus* J1074 and *Streptomyces lividans* TK24 were used as hosts for gene heterologous expression. All *E. coli* strains were grown in Luria-Bertani broth [17]. *Streptomyces* were grown on manni-

tol/soy (MS) agar plates for spores [15]. For STs production, *S. coelicolor* CIM1001 was cultured in Medium D (soybean meal 2%, yeast extract 0.5%, starch 2.5%, $CaCO_3$ 0.35%, $CuSO_4$ 0.007%, $FeSO_4$ 0.001%, $MnCl_2$ 0.008%, $ZnSO_4$ 0.002%, pH7.4) at 28°C, 220 r min^{-1}. The concentrations of antibiotics used for selection were ampicillin 100 μg mL^{-1}, kanamycin 50 μg mL^{-1}, apramycin 50 μg mL^{-1}, chloroamphenicol 25 μg mL^{-1} and thiostrepton 10 μg mL^{-1}.

Plasmid pBluscript II SK(−) was from a commercial source (Stratagene, La Jolla, CA, USA), pUC119::kan was described previously [18], and pUWL201PW was kindly provided by Dr. Udo Wehmeier [19]. Plasmid pUW201PW-oriT was constructed by inserting an *oriT* cassette into the *Kpn* I site of pUWL201PW. The *oriT* cassette was cloned from pSET152 by PCR using primers oriTF (5′-GGGGTACCTGGTTGGCTTGGTTTCATC-3′) and oriTR (5′-GGGGTACCTGCCAAAGGGTTCGTGTAG-3′).

1.2 DNA manipulation

Routine DNA manipulation was carried out according to standard methods [17]. Gene disruption and complementation were as described previously [20–22]. *E. coli-Streptomyces* conjugation and λ RED-mediated recombination were performed as described previously [15,16]. PCR reactions were performed with Taq DNA polymerase (Takara) or KOD-Plus DNA polymerase (Toyobo) according to the manufacturers' instructions.

1.3 Sequencing and bioinformatics analyses

Whole genome Solexa sequencing of *S.* sp. TP-A0356 was described previously [23], and the sequence of the ST cluster is available in GenBank under accession number KC935381. FramePlot 4.0 beta (http://nocardia.nih.go.jp/fp4) was used to deduce possible open reading frames (ORFs). A BLAST search was used to predict gene functions. Alignments were performed with CLUSTALW. Prediction of the NRPS A domain substrates was performed using the NRPS predictor 2 software package and the PKS/NRPS Analysis web site.

1.4 Construction of ST heterologous expression strains

Construction of the cosmid library of *S.* sp. TP-A0356 genome was described previously [23]. The cosmids containing putative genes responsible for the biosynthesis of streptothricins were screened by PCR using primers *stnD*-For (5′-CGACAGCGGCTGGGACTTCC-3′) and *stnD*-Rev (5′-ACCCAGATGAAGGACGCCCG-3′), *stnO*-For (5′-CTGGACGAGGGTTTCTACGG-3′) and *stnO*-Rev (5′-GCACAGGTCGAGCAGGTGG-3′). Cosmid pTG6001 was introduced into *S. coelicolor* M145, *S. albus* J1074 and *S. lividans* TK24 by *E. coli-Streptomyces* conjugation to generate the strains CIM1001, CIM1002 and CIM1003, respectively.

1.5 Construction and complementation of the Δ*stnO* mutant CIM1004

The kan^R cassette was cloned from pUC119::kan with primers *stnO*-DF (5′-GCGATACGACCGAAACCGATGGGATTGGGTGGCGGGTTGGTGCCCCTGGATACCGCTCG-3′) and *stnO*-DR (5′-GGTTGTGCGACTGTCCCTTCTATCCCTGACCGGTGCCCTGATGAACCCCAGAGTCCCGC-3′). The amplified 0.98-kb fragment was then used to replace *stnO* on pTG6001 via λ RED-mediated recombination to generate pCIM1001 (Figure S2). Successful construction of pCIM1001 was verified by restriction enzymes digestion and DNA sequencing. Introduction of pCIM1001 into *S. coelicolor* M145 by *E. coli-Streptomyces* conjugation afforded the Δ*stnO* mutant strain CIM1004.

To complement CIM1004, the *stnO* gene was amplified from pTG6001 using primers *stnO*-EF (5′-GGAATTCCATATGGACACAGCACGCAGCCAG-3′) and *stnO*-ER (5′-GAAGATCTCAGCGGACGCAGACTGCTCAT-3′). The resulting 1.4-kb fragment was then inserted into the *Eco*R V site of pBluscript II SK(−) for sequencing. Correct *stnO* was excised as a 1.4-kb *Nde* I/*Bgl* II fragment and inserted into the *Nde* I/*Bam*H I sites of pUWL201PW-oriT to generate pCIM1002. Introduction of pCIM1002 into CIM1004 afforded the Δ*stnO* complemented strain CIM1005.

1.6 Isolation of streptothrisamine

S. coelicolor CIM1004 was cultured in medium D for 5 d at 28°C, 220 r min^{-1}. The supernatant of 2 L culture broth was obtained by centrifugation and extracted with 2 L chloroform. The aqueous layer from centrifugation was mixed with 200 g Diaion HP-20 resin. After removing resin by filtration, the filtrate was loaded on a Dowex 50W-X2 column (100–200 mesh, NH_4^+ form, 200 g). After washed with 300 mL distilled water, the sample was eluted stepwise with 300 mL of 1.0 and 2.0 mol L^{-1} NH_4HCO_3. The 1.0 mol L^{-1} NH_4HCO_3 fraction containing streptothrisamine was desalted on a Sephadex LH-20 column (3.0 cm×90 cm). The water eluted fraction was then lyophilized to get a white powder. After dissolved in a small volume of water, streptothrisamine was refined by preparative high-performance liquid chromatography (HPLC; Zorbas SB-Aq PrepHT, 7 μm, 21.2 mm×250 mm, Agilent, Santa Clara, CA, USA) at a flow rate of 7 mL min^{-1} with 2% acetonitrile/water in 0.05% n-heptafluorobutyric acid and 0.05% formic acid. The purified streptothrisamine was lyophilized to give a white powder and used for NMR analysis.

1.7 Antibacterial assays

Staphylococcus aureus and *Bacillus subtilis* were used as indicator stains for antibacterial bioactivity tests of STs. The bacteria were cultured in LB medium at 37°C and 220 r min^{-1} overnight. Then, 1 mL of the broth was added to 200 mL LB agar medium (melted and kept at 50°C), and the mixture was laid on petri dishes. Aliquots of 100 μL fermentation broth were used for every test. Inhibition zones were observed after incubating the plates at 37°C for 12 h.

1.8 Spectroscopic analysis of STs

For STs production, strains for gene heterologous expression were cultured in different media (including medium D) at 28°C and 220 r min^{-1}. After 5 d fermentation, the broth was harvested by centrifugation (6000 r min^{-1}, 5 min), and then subjected to the bioassay test against *S. aureus* and *B. subtilis*. The broth with inhibitory activity was filtered through a 0.22 μm Millipore membrane and subjected to HPLC analysis.

HPLC analysis was carried out with an Apollo C18 column (5 μm, 4.6 mm×250 mm, Alltech, Deerfield, IL, USA) on a Shimadzu HPLC system (Shimadzu, Kyoto, Japan). The column was developed with a 20 min linear gradient using solution A (10 mmol L^{-1} 1-heptane sulfonic acid sodium salt solution with 0.072% acetic acid) and acetonitrile at a flow rate of 1.0 mL min^{-1}. Percentage of acetonitrile was changed linearly from 5% to 35% at 0–10 min, and from 35% to 50% at 10–20 min. The detection wavelength was 190 nm. ST standard was purchased from Santa Cruz Biotechnology, Inc. (Santa Cruz, CA, USA).

For liquid chromatography coupled with mass spectrometry (LC-MS) analysis, an Agilent 1260/6460 Triple-Quadrupole LC/MS system was used. The spray voltage was set to 4.0 kV and heated transfer capillary temperature was set to 350°C. High resolution mass spectrometry (HR-MS) analysis was carried out with Aligent 1200 HPLC system and 6520 QTF-MS system (Agilent, Santa Clara, CA, USA) with the electrospray ionization (ESI) source. NMR analysis was carried out in D_2O with a 400 mol L^{-1} Hz Bruker NMR Spectrometer (Bruker, Billerica, MA, USA).

2 Results

2.1 Uncover a cryptic ST cluster in *S.* sp. TP-A0356 genome

S. sp. TP-A0356 was sequenced previously to clone the yatakemycin cluster [23]. When its genome was analyzed using ST biosynthetic genes as probes, one putative ST biosynthetic gene cluster (GenBank accession No. KC935381) was uncovered (Figure 1B). Comparative genomic analysis revealed that the organization of the ST cluster from *S.* sp. TP-A0356 is very similar to the other five ST gene clusters (Figure S1). Moreover, all ST related genes in the reported ST cluster from *S. rochei* NBRC 12908 could be found in KC935381, indicating that a complete ST cluster is embedded in *S.* sp. TP-A0356 genome. The major difference between the ST cluster from *S.* sp. TP-A0356 and that from *S. rochei* NBRC12908 is that there are two extra MerR type regulators located at the left terminus of the former cluster (Figure 1B), while no regulatory gene can be found in the latter one.

2.2 Identification of the ST cluster by heterologous expression

To study the cryptic ST cluster from *S.* sp. TP-A0356, a constructed cosmid library of this strain's genome was screened by PCR using *stnD*-For/*stnD*-Rev and *stnO*-For/*stnO*-Rev as primers. A cosmid, pTG6001 was selected and used in subsequent heterologous expression. The 41-kb fragment in pTG6001 contains all ST biosynthesis related genes (from *orf-3* to *stnU*).

Cosmid pTG6001 was introduced into *S. coelicolor* M145, *S. albus* J1074 and *S. lividans* TK24 by *E. coli-Streptomyces* conjugation to generate strains CIM1001, CIM1002 and CIM1003, respectively. The three recombinant strains were incubated in different media and screened for STs production by antibacterial assays against *S. aureus* and *B. subtilis*. Positive samples were then analyzed by HPLC. When cultured in medium D, CIM1001 broth showed high antibacterial activity. HPLC detection revealed two peaks corresponding to ST-F (**1**) and ST-D (**2**) (Figure 2). The identities of ST-F and ST-D were further confirmed by co-injection with standards, MS data ($[M+H]^+$ ions at *m/z* 503.1 for ST-F and $[M+H]^+$ ions at *m/z* 759.2 for ST-D), and tandem mass spectrometry (MS/MS) analysis (Figure S3).

Besides the ST-F and ST-D peaks, two new peaks at 16.5 min and 16.9 min were observed in CIM1001 culture. Both compounds (**3** and **4**) showed similar UV-spectra as STs, implying that they were ST analogs. The molecular weights of compounds **3** and **4** were determined to be 1056 Da ($[M+H]^+$ ions at *m/z* 1057.4) and 1184 Da ($[M+H]^+$ ions at *m/z* 1185.5), respectively. High resolution ESI-MS of **3** yielded $[M+H]^+$ and $[M+Na]^+$ ions at *m/z* 1057.6481 and 1079.6296, showed that the molecular formula is $C_{45}H_{84}N_{16}O_{13}$ (1057.6482 and 1079.6301 calculated for the $[M+H]^+$ and $[M+Na]^+$ ions, respectively). High resolution ESI-MS of compound **4** proposed a molecular formula $C_{51}H_{96}N_{18}O_{14}$ with $[M+H]^+$ ion at *m/z* 1185.7432 (1185.7432 calculated for the $[M+H]^+$ ion). MS analysis suggested that compound **3** is acetylated streptothricin B (Ac-ST-B) and compound **4** is acetylated streptothricin A (Ac-ST-A). MS/MS analysis of both compounds also veri-

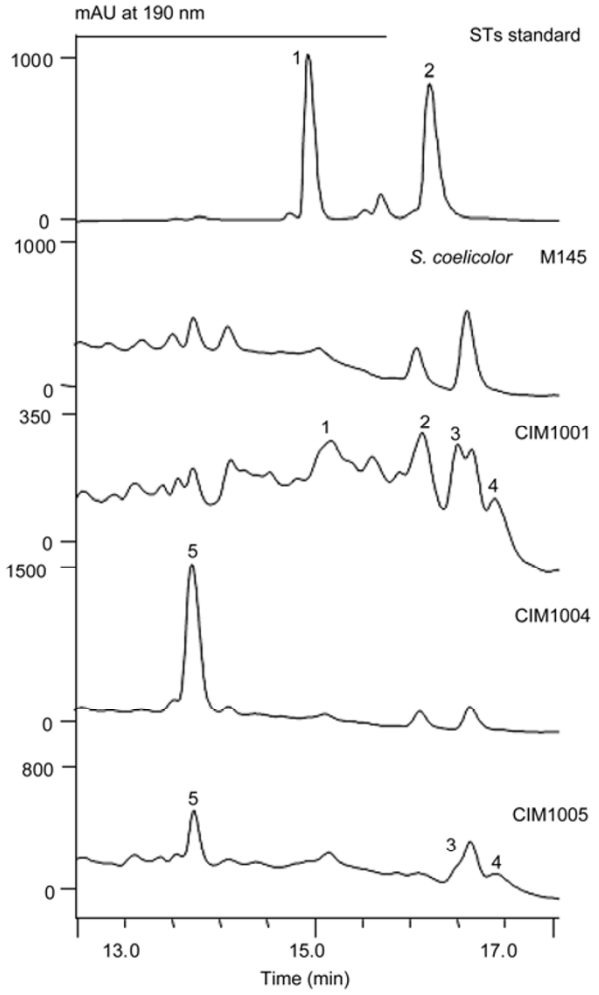

Figure 2 HPLC analysis of streptothricins production in different *Streptomyces* strains. The host strain *S. coelicolor* M145 was used as a control. CIM1001 is the *S. coelicolor* strain containing the whole ST biosynthetic gene cluster; CIM1004 is the Δ*stnO* mutant; and CIM1005 is the *stnO* in trans complemented strain.

fied the structure assignment (Figure S4). Acetylation is a well-studied resistance mechanism of STs, and the acetyl group has always been added at β-amino group of the β-Lys directly bonded to sugar moiety. Fragmentation patterns of Ac-ST-A strongly suggested that the acetylation took place at the first β-Lys (Figure S4A).

2.3 Functional analyses of individual ST genes

Production of STs in CIM1001 proved that pTG6001 contains a complete ST biosynthetic gene cluster. Twenty-four genes within the cluster (spanning 30.5 kb) are proposed to be involved in ST biosynthesis. Based on previous investigations on ST biosynthesis and bioinformatics analyses, functions of most genes are proposed (Table 1).

Four genes (*stnE*, *stnO*, *stnR* and *stnS*) are proposed to be involved in the poly-β-Lys chain biosynthesis. Both StnE and StnS are stand-alone NRPS A domains. StnE shows 87.5% identity to Orf5 from *S. rochei* NBRC12908 ST cluster (GenBank accession No. AB684619), which activates one β-Lys and loads it onto the phosphopantetheine arm of the corresponding PCP domain [11]. StnS should be responsible for the poly β-Lys chain elongation, since it shares 77.1% identity with Orf19 from AB684619. StnR is an NRPS PCP-C didomain protein showing 62.8% identity with Orf18 from AB684619 [11]. StnO is proposed to be an aminomutase showing 68% identity to CmnP in capreomycin biosynthesis from *Saccharothrix mutabilis*, which catalyzes the conversion from L-Lys to (3*S*)-β-Lys [24].

Four genes related to the sugar moiety biosynthesis are *stnG*, *stnH*, *stnI* and *stnQ*. StnI shows significant similarity to several LmbE like proteins (e.g., 20.9% identity to BC1534 from *Bacillus cereus*), which can remove the acetyl group of N-acetyl glucosamine (GlcNAc) from a bunch of different substrates [25]. StnH and StnQ show high similarities to putative carbamoyl-phosphate synthetase ABB37219 (34% identity) from *Desulfovibrio alaskensis* G20 and carbamoyl transferase TobZ (53% identity) in tobramycin biosynthesis from *Streptoalloteichus tenebrarius* [26], respectively. StnG is a putative glycosyl transferse similar to YP_006230477 (20.9% identity) from *Bacillus* strain [27].

Eight genes (*stnC*, *stnD*, *stnF*, *stnJ*, *stnK*, *stnL*, *stnM* and *stnN*) may be involved in streptolidine lactam formation. StnC shows high similarity with a thioesterase CinE (42% identity) from *S.* sp JS360 [28]. Inactivation of *sttC*, the counterpart of *stnC* in *S. rochei* F20, dramatically decreased STs production [9]. StnD shows significant similarity with a macrolide phosphotransferase (43% identity) from *Saccharopolyspora erythraea* [29]. Inactivation of its homolog gene *sttD* blocked STs production in *S. rochei* F20 [9], implying that it was necessary for STs biosynthesis. StnF is a homolog of peptidase/esterase with 36% identity to cocaine esterase CocE from *Rhodococcus* sp. MB1 [30]. StnJ is a didomain protein with an N-terminal NAD-dependent oxidoreductase domain (Pfam01408) and a C-terminal NAD dependent epimerase/dehydratase domain (Pfam01370). Gene *stnK* encodes a protein showing 33% identity to SxtR, an acyltransferase from saxitoxin biosynthetic cluster in cyanobacterium *Cylindrospermopsis raciborskii* T3 [31]. StnM is an NRPS A-PCP didomain protein. Bioinformatics analysis of the StnM A domain predicted that it may activate Arg, Trp or 3-amino-6-hydroxy-2-piperidone. StnL and StnN are homologs of VioC (46% identity) and VioD (40% identity) from viomycin biosynthesis, which catalyze the α-ketoglutarate dependent β-hydroxylation of L-Arg and subsequent pyridoxal phosphate dependent cyclization to form capreomycidine [32].

Three genes are proposed to be involved in the self-protection of ST producer including a homolog of the well conserved ST acetyl transferase gene *stnA* and two ABC transporter genes *stnT* and *stnU*.

Regulatory genes *orf-2* and *orf-3* encode two proteins sharing high similarity with MerR family transcriptional

Table 1 Proposed functions of ORFs in the Streptothricin biosynthetic gene cluster from *S.* sp. TP-A0356

Gene	Size (amino acids)	Putative function	Homolog (GenBank Ac. No.)	% Similarity/ % identity
orf-3	128	Regulatory protein	Regulator (AEM85931)	94/89
orf-2	124	Regulatory protein	Regulator (EFE74449)	76/59
orf-1	300	Dehydrogenase	Dehydrogenase (EEM84513)	72/50
stnA	195	Acetyltransferase	Nat1 (CAA51674)	87/81
stnB	218	Unknown	Hypothetical protein (ACU74814)	57/46
stnC	263	Thiosterase	CinE (CBW54675)	53/42
stnD	311	Phosphotransferase	MphB (CAL99425)	55/43
stnE	505	NRPS (A)	NpsA (CAC86008)	91/88
stnF	558	Peptidase/esterase	CocE (Q9L9D7)	50/36
stnG	387	Glycosyl transferase	Glycosyl-transferase (EEL37055)	58/37
stnH	410	Carbamoyl phosphate synthetase	Phosphoribosylglycinamide synthetase (ABB37219)	50/34
stnI	244	Deacetylase	LmbE (EEM98880)	47/32
stnJ	674	Dehydrogenase	oxidoreductase domain protein (HP35370)	44/31
		dehydratase	NAD-dependent epimerase or dehydratase (ACL44529)	51/33
stnK	247	Transferase	SxtR (ABI75112)	54/33
stnL	344	Oxygenase	VioC (AAP92493)	61/46
stnM	658	NRPS (A-PCP)	PvdI (ABA73588)	51/34
stnN	382	Aminotransferase	VioD (AAP92494)	51/40
stnO	467	Aminomutase	CmnP (ABR67759)	79/68
stnP	225	Unknown	WD-40 repeat-containing protein (AEW98573)	77/65
stnQ	590	Carbamoyltransferase	TobZ (CAH18554)	68/53
stnR	597	NRPS (C-PCP)	NpsB (CBC86013)	85/82
stnS	510	NRPS (A)	NpsA (CBC86008)	62/55
stnT	564	Transporter	StrV(CAH94305)	57/43
stnU	482	Transporter	GetC(CBL93714)	67/48

regulators (Orf-2 shows 62% identity with ZP_21027493 [33], and Orf-3 shows 89% identity with YP_004816211). Preliminary gene inactivation studies showed that Orf-2 could significantly influence STs production (unpublished data).

Three remaining genes in the cluster are *stnB*, *stnP* and *orf-1*. Hypothetical protein StnB shows no significant similarity to any known protein. StnP shows 65% identity to a WD-40 repeat containing protein from *S. cattleya*, which may have regulatory functions [34]. Orf-1 is homolog of a putative dehydrogenase (50% identity) from *Bacillus thuringiensis*. It only exists in the cluster from *S.* sp. TP-A0356 and has no homologs in the other ST clusters, indicating it may not be a ST biosynthetic gene.

2.4 Evaluation of the Δ*stnO* mutant

StnO was proposed to be an aminomutase converting L-Lys to (3*S*)-*β*-Lys, which was used as precursors to build the poly-*β*-Lys chain. Gene *stnO* in pTG6001 was replaced by a kanamycin resistance cassette via *λ*-Red mediated PCR-targeting strategy to generate pCIM1001. The Δ*stnO* mutant CIM1004 was obtained by introducing pCIM1001 into *S. coelicolor* M145. When cultured in medium D, productions of ST-F, ST-D, Ac-ST-A and Ac-ST-B were all abolished in CIM1004. A new peak at 13.7 min was observed in HPLC trace, which had a UV spectrum very similar to the ST analogs (Figure 2). ESI-MS showed the molecular weight of this compound was 374 Da ($[M+H]^+$ ion at *m/z* 375.1). High resolution ESI-MS determined its molecular formula as $C_{13}H_{22}N_6O_7$ by an $[M+H]^+$ ion at *m/z* 375.1624 (calculated 375.1623). The compound was finally determined to be streptothrisamine by MS/MS and ^{1}H NMR data (Figures 1 and S5, Table S1).

Complementation of CIM1004 was carried out by constitutively expressing *stnO* in trans to generate CIM1005. The production of streptothrisamine decreased dramatically in CIM1005, and production of Ac-ST-B and Ac-ST-A were observed (Figure 2). Although the complementation did not fully restore STs production in CIM1005, it proved that transcription of the genes downstream *stnO* was not blocked by *stnO* inactivation.

3 Discussion

Along with the explosion of the number of sequenced microbial genomes, more and more cryptic secondary clusters have been found. Genome mining is becoming a powerful approach to uncover natural products with novel structures and to study secondary metabolites biosynthesis [35,36]. Considering the wide distribution of STs in nature, we assumed there are some ST clusters embedded in the sequenced genomes. Database searching indeed discovered two putative ST clusters in *S.* sp. C and *S.* sp. TP-A0356 genomes. The ST cluster from *S.* sp. TP-A0356 was confirmed finally by its successful heterologous expression in *S. coelicolor* M145 and

the subsequent gene inactivation experiment.

To our surprise, we found that two new ST analogs Ac-ST-A and Ac-ST-B were produced besides ST-F and ST-D when the ST cluster from *S.* sp. TP-A0356 was expressed in *S. coelicolor*. It is likely that Ac-ST-A and Ac-ST-B are derived from ST-A and ST-B by acetylating at β-amino group of the β-Lys directly bonded to streptothrisamine. Bioactivity tests showed that both Ac-ST-A and Ac-ST-B possessed very low antibacterial activities (unpublished data), which is in accordance with the fact that β-Lys acetylation is a typical detoxification strategy of STs [37]. Since no ST acetyl transferase encoding gene was found in *S. coelicolor* M145 genome, the acetyl group may be added by the ST acetyl transferase StnA.

Based on information from bioinformatics analyses and previous results, a putative biosynthetic pathway for STs was proposed (Figure 3). In viomycin biosynthesis, L-Arg is converted to (2*S*,3*R*)-capreomycidine by VioC and VioD through β-hydroxylation and subsequent dehydration dependent cyclization [32]. A similar pathway may exist in ST biosynthesis since StnL and StnN show significant similarities with VioC and VioD, respectively. However, feeding experiments failed to incorporate (2*S*,3*R*)-capreomycidine into ST [38], indicating a tethered biosynthetic mechanism might be involved. Discovery of StnM, an A-PCP didomain protein possibly activating L-Arg analogs in the ST cluster supported this hypothesis. It was proposed that StnL, StnM and StnN worked together to form a PCP tethered capreomycidine from L-Arg. Capreomycidine was then converted to streptolidine through an unknown mechanism. The tethered intermediate must be hydrolyzed (possibly by StnC) before the lactam ring formation, since streptolidine was proved to be an intermediate of ST [38]. The conjugation of streptolidine lactam and sugar moiety is reminiscent of paromamine and neamine biosynthesis in aminoglycoside pathways, which also form a glycosidic bond with N-acetyl glucosamine [39]. StnG was proposed to catalyze the glycosylation of streptolidine, and StnH and StnQ should be responsible for the carbamoyl group addition. Streptothrisamine was generated by a de-acetylation step catalyzed by StnI. Feeding experiments have shown already that β-lysines in ST are generated from L-Lys [8]. Accumulation of streptothrisamines in the Δ*stnO* mutant CIM1004 proved that StnO is the aminomutase responsible for this conversion. The β-lysines are then adenylated by the stand-alone A domain proteins StnE and StnS, tether to the PCP domain of StnR and form the poly-β-Lys chain. The C domain of StnR attaches the poly-β-Lys chain to streptothrisamine finally to form mature STs [11].

Conclusively, we identified a ST biosynthetic gene cluster from the genome of yatakemycin producer *S.* sp. TP-A0356. Two new ST analogs (Ac-ST-A and Ac-ST-B) together with ST-F, ST-D were produced when this ST cluster was heterologously expressed in *S. coelicolor* M145. Gene inactivation results suggested that the aminomutase StnO was responsible for β-Lys formation, which filled the last gap of poly-β-Lys chain biosynthesis. Biosynthesis of the unique streptolidine moiety and the mechanism of sugar incorporation are still mysteries to be solved.

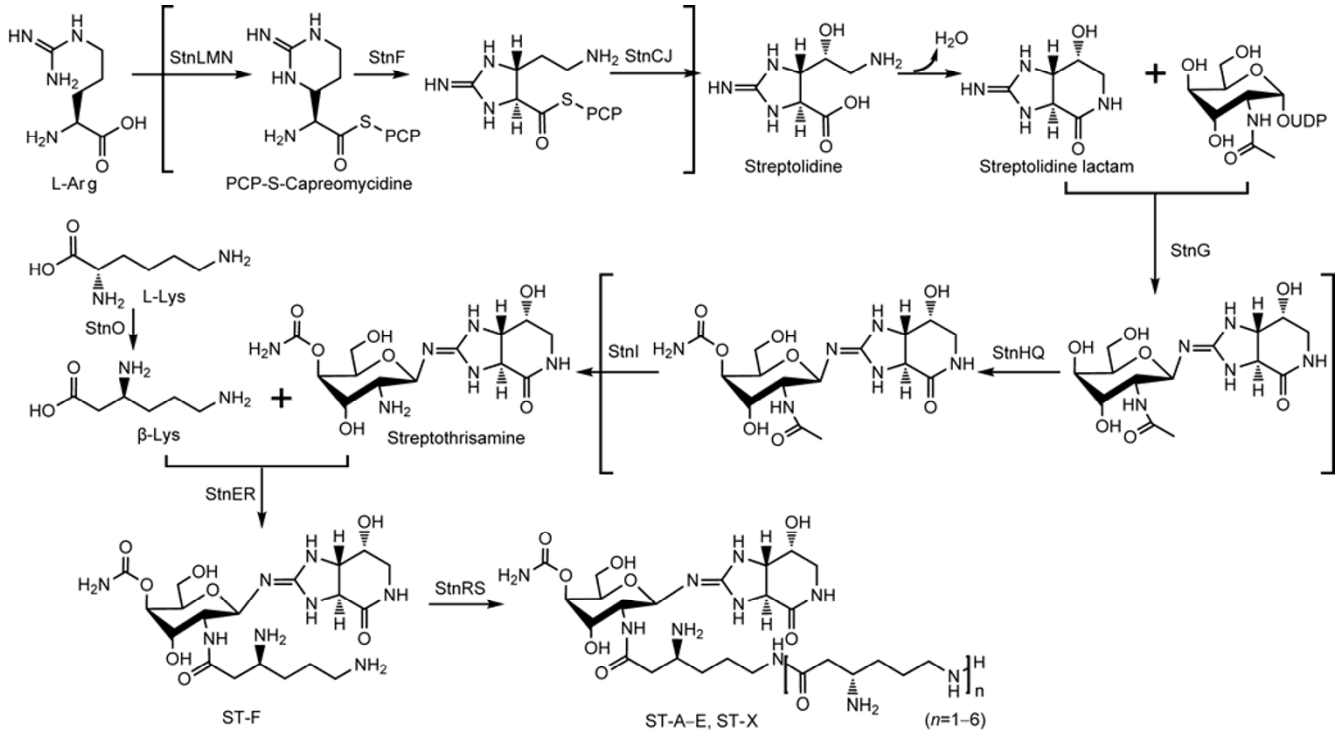

Figure 3 Proposed biosynthetic pathway of streptothricins. Postulated intermediates in STs biosynthesis are bracketed.

We thank Prof. Yasuhiro Igarashi (Toyama Prefectural University, Japan) for providing Streptomyces sp. TP-A0356 and the Analytical Instrumentation Center of State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences for supports in MS analyses. This work was supported in part by the National Natural Science Foundation of China (31170037), Ministry of Science and Technology of China (2013CB734003) and the China Postdoctoral Science Foundation (2013M530755).

1 Waksman S A. Production and activity of streptothricin. J Bacteriol, 1943, 46: 299–310
2 Romer W, Hesse G, Miosga N, et al. Chemical determination of the streptothricin antibiotic nourseothricin. Arch Exp Vet Med, 1986, 40: 693–698
3 Ohba K, Nakayama H, Furihata K, et al. Albothricin, a new streptothricin antibiotic. J Antibiot, 1986, 39: 872–875
4 Borders D B, Sax K J, Lancaste J E, et al. Structures of LL-AC541 and LL-AB664: new streptothricin-type antibiotics. Tetrahedron, 1970, 26: 3123–3133
5 Inamori Y, Amino H, Tsuboi M, et al. Biological-activities of racemomycin-B, β-lysine rich streptothricin antibiotic, the main component of *Streptomyces lavendulae* Op-2. Chem Pharm Bull, 1990, 38: 2296–2298
6 Witte W. Selective pressure by antibiotic use in livestock. Int J Antimicrob Agents, 2000, 16: S19–S24
7 Jelenska J, Tietze E, Tempe J, et al. Streptothricin resistance as a novel selectable marker for transgenic plant cells. Plant Cell Rep, 2000, 19: 298–303
8 Thiruvengadam T K, Gould S J, Aberhart D J, et al. Biosynthesis of streptothricin-F. 5. Formation of β-lysine by *Streptomyces* L-1689-23. J Am Chem Soc, 1983, 105: 5470–5476
9 Fernández-Moreno M A, Vallin C, Malpartida F. Streptothricin biosynthesis is catalyzed by enzymes related to nonribosomal peptide bond formation. J Bacteriol, 1997, 179: 6929–6936
10 Grammel N, Pankevych K, Demydchuk J, et al. A β-lysine adenylating enzyme and a β-lysine binding protein involved in poly β-lysine chain assembly in nourseothricin synthesis in *Streptomyces noursei*. Eur J Biochem, 2002, 269: 347–357
11 Maruyama C, Toyoda J, Kato Y, et al. A stand-alone adenylation domain forms amide bonds in streptothricin biosynthesis. Nat Chem Biol, 2012, 8: 791–797
12 Gould S J, Lee J N, Wityak J. Biosynthesis of streptothricin-F. 7. The fate of the arginine hydrogens. Bioorg Chem, 1991, 19: 333–350
13 Palaniswamy V A, Gould S J. Biosynthesis of streptothricin-F. 6. Formation and intermediacy of D-glucosamine in *Streptomyces* L-1689-23. J Chem Soc, Perkin Trans 1, 1988, 8: 2283–2286
14 Baltz R H. Antimicrobials from actinomycetes: back to the future. Microbe, 2007, 2: 125–131
15 Kieser T, Bibb M J, Buttner M J, et al. Practical *Streptomyces* Genetics. Norwich: John Innes Foundation, 2000
16 Gust B, Challis G L, Fowler K, et al. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc Natl Acad Sci USA, 2003, 100: 1541–1546
17 Sambrook J, Fritsch T, Maniatis E F. Molecular Cloning: A Laboratory Manual. New York: Cold Spring Harbor Laboratory Press, 2001
18 Liu G, Tian Y, Yang H, et al. A pathway-specific transcriptional regulatory gene for nikkomycin biosynthesis in *Streptomyces ansochromogenes* that also influences colony development. Mol Microbiol, 2005, 55: 1855–1866
19 Doumith M, Weingarten P, Wehmeier U F, et al. Analysis of genes involved in 6-deoxyhexose biosynthesis and transfer in *Saccharopolyspora erythraea*. Mol Gen Genet, 2000, 264: 477–485
20 Pan Y, Liu G, Yang H, et al. The pleiotropic regulator AdpA-L directly controls the pathway-specific activator a of nikkomycin biosynthesis in *Streptomyces ansochromogenes*. Mol Microbiol, 2009, 72: 710–723
21 Li R, Liu G, Xie Z, et al. PolY, a transcriptional regulator with ATPase activity, directly activates transcription of *polR* in polyoxin biosynthesis in *Streptomyces cacaoi*. Mol Microbiol, 2010, 75: 349–364
22 Xu G, Wang J, Wang L, et al. Pseudo gamma-butyrolactone receptors respond to antibiotic signals to coordinate antibiotics biosynthesis. J Biol Chem, 2010, 285: 27440–27448
23 Huang W, Xu H, Li Y, et al. Characterization of yatakemycin gene cluster revealing a radical S-adenosylmethionine dependent methyltransferase and highlighting spirocyclopropane biosynthesis. J Am Chem Soc, 2012, 134: 8831–8840
24 Felnagle E A, Rondon M R, Berti A D, et al. Identification of the biosynthetic gene cluster and an additional gene for resistance to the antituberculosis drug capreomycin. Appl Environ Microbiol, 2007, 73: 4162–4170
25 Deli A, Koutsioulis D, Fadouloglou V E, et al. LmbE proteins from *Bacillus cereus* are de-N-acetylases with broad substrate specificity and are highly similar to proteins in *Bacillus anthracis*. FEBS J, 2010, 277: 2740–2753
26 Parthier C, Gorlich S, Jaenecke F, et al. The O-carbamoyltransferase TobZ catalyzes an ancient enzymatic reaction. Angew Chem Int Ed, 2012, 51: 4046–4052
27 Song J Y, Kim H A, Kim J S, et al. Genome sequence of the plant growth-promoting *rhizobacterium Bacillus* sp. strain JS. J Bacteriol, 2012, 194: 3760–3761
28 Rachid S, Huo L, Herrmann J, et al. Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. Chembiochem, 2011, 12: 922–931
29 Oliynyk M, Samborskyy M, Lester J B, et al. Complete genome sequence of the erythromycin-producing bacterium *Saccharopolyspora erythraea* NRRL 23338. Nat Biotechnol, 2007, 25: 447–453
30 Bresler M M, Rosser S J, Basran A, et al. Gene cloning and nucleotide sequencing and properties of a cocaine esterase from *Rhodococcus* sp. strain MB1. Appl Environ Microbiol, 2000, 66: 904–908
31 Moustafa A, Loram J E, Hackett J D, et al. Origin of saxitoxin biosynthetic genes in cyanobacteria. PLoS ONE, 2009, 4: e5758
32 Ju J, Ozanick S G, Shen B, et al. Conversion of (2*S*)-arginine to (2*S*,3*R*)-capreomycidine by VioC and VioD from the viomycin biosynthetic pathway of *Streptomyces* sp. strain ATCC11861. Chembiochem, 2004, 5: 1281–1285
33 Alonso-Vega P, Normand P, Bacigalupe R, et al. Genome sequence of *Micromonospora lupini* Lupac 08, isolated from root nodules of *Lupinus angustifolius*. J Bacteriol, 2012, 194: 4135
34 van der Voorn L, Ploegh H L. The WD-40 repeat. FEBS Lett, 1992, 307: 131–134
35 Winter J M, Behnken S, Hertweck C. Genomics-inspired discovery of natural products. Curr Opin Chem Biol, 2011, 15: 22–31
36 Challis G L. Mining microbial genomes for new natural products and biosynthetic pathways. Microbiology, 2008, 154: 1555–1569
37 Krugel H, Fiedler G, Haupt I, et al. Analysis of the nourseothricin-resistance gene (*nat*) of *Streptomyces noursei*. Gene, 1988, 62: 209–217
38 Jackson M D, Gould S J, Zabriskie T M. Studies on the formation and incorporation of streptolidine in the biosynthesis of the peptidyl nucleoside antibiotic streptothricin F. J Org Chem, 2002, 67: 2934–2941
39 Llewellyn N M, Spencer J B. Biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. Nat Prod Rep, 2006, 23: 864–874

Supporting Information

Figure S1 Comparative analysis of six streptothricin biosynthetic gene clusters from different *Streptomyces* strains. The two clusters at the bottom of the figure were uncovered in this study by genome mining using specific ST genes as probes.

Figure S2 Construction of gene *stnO* inactivated cosmid pCIM1001. The whole *stnO* gene was replaced by a kanamycin resistance cassette.

Figure S3 MS/MS analysis of streptothricins F and D. A, ST-F. B, ST-D.

Figure S4 MS/MS analysis of the streptothricins identified in this study. A, Ac-ST-A. B, Ac-ST-B. High resolution MS spectra for both compounds are also presented.

Figure S5 Structure elucidation of streptothrisamine. The structures of streptothrisamine (A) and 10-decarbamoyl-12-carbamoyl streptothrisamine (B) are shown. C and D, MS/MS (C) and high resolution MS data (D) of streptothrisamines. E, ^{1}H NMR spectrum of streptothrisamines. Along with streptothrisamine, some amount of 10-decarbamoyl-12-carbamoyl streptothrisamine was also detected by ^{1}H NMR in *S. coelicolor* CIM1004.

Table S1 ^{1}H NMR (400 MHz) data for streptothrisamine and 10-decarbamoyl-12-carbamoyl streptothrisamine

The supporting information is available online at life.scichina.com and www.springerlink.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.